DOCUMENT RESUME

ED 130 839 SE 021 375

AUTHOR Hayward, Robert R.
TITLE The Developing and Field Testing of an Instrument Using the Planetarium to Evaluate the Attainment of the Concept of Annual Motion.
PUB DATE Apr 76
NOTE 17p.; Paper presented at the annual meeting of the National Association for Research in Science Teaching (49th, San Francisco, California, April 23-25, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Astronomy; Criterion Referenced Tests; *Educational Research; Evaluation; *Planetariums; Science Education; Secondary Education; *Secondary School Science; Space Sciences; *Test Construction; Tests
IDENTIFIERS Research Reports

ABSTRACT

The major problem of this study was the development of an instrument, the Planetarium Test on Annual Motion (PTAM), that would use an observational format in the planetarium as a reasonable substitute for the real sky to evaluate the attainment of instructional objectives related to selected concepts on the annual motions of the sun, moon, and planets. The instrument was designed to evaluate student performance on 17 objectives, 3 processes of science, and 3 intellectual skills. Thirteen concepts on annual motion were developed. Seventeen behavioral objectives congruent with the concepts were developed and two items written for each objective. The 34-item instrument was field tested with 471 sixth-grade students who received one of three treatments: a planetarium program, a similar classroom program, and no instruction (control). Both the planetarium and classroom treatments produced general gains in achievement more than mastery of concepts. The planetarium was significantly superior to the classroom on many mean and criterion scores for content objectives and, in general, the planetarium instruction fared well on objectives requiring evaluation of observational relationships and higher skills using these observations. (Author/MH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED130839

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE DEVELOPING AND FIELD TESTING OF AN INSTRUMENT USING THE PLANETARIUM TO EVALUATE THE ATTAINMENT OF THE CONCEPT OF ANNUAL MOTION

by
Robert R. Hayward
Fernbank Science Center
Atlanta, Georgia 30307

A Paper Presented to
The 49th Annual Meeting of
The National Association for Research in Science Teaching
San Francisco, California
April 23, 1976

SE 021 375

THE DEVELOPING AND FIELD TESTING OF AN INSTRUMENT USING THE PLANETARIUM TO EVALUATE THE ATTAINMENT OF THE CONCEPT OF ANNUAL MOTION

Introduction

The major problem of this study was the development of a testing instrument, the Planetarium Test on Annual Motions (PTAM), that uses an observational format in the planetarium as a reasonable substitute for the real sky to evaluate the attainment of instructional objectives related to selected concepts of the annual motions of the sun, moon, and planets. The instrument was designed to evaluate student performance on seventeen content objectives, three processes of science, and three intellectual skills. It also was designed to be applicable to the evaluation of any means of instruction whose instructional objectives are the same as those evaluated by the PTAM.

Since the field of planetarium education is of reasonably recent vintage, there are numerous research questions that need to be answered to further guide the direction of education in the field. The effectiveness of the planetarium as a teaching institution has been evaluated to a limited extent, usually vis á vis another method such as a classroom presentation. However, such evaluations have been few and usually have been limited to the investigator's special problem. The effectiveness of the planetarium as a teaching tool remains basically in the realm

of the subjective feelings and expressions of those working in the field.

Previous evaluation instruments in the field have usually consisted of paper and pencil tests. However, paper and pencil tests tend to show what a student can do on paper with a pencil. Warneking (1970) put forth a caution in a call for evaluation of planetarium education efforts. He said "We should be aware of the 'cue' theory when testing recall of factual material. This states that a class tested for recall in the same surroundings as that in which the instruction was initially received will do better than one tested in a different setting" (p. 15). In 1972 Dean and Lauck reinforced the point and noted that one possible explanation for the inconsistent results among these studies was that the wrong kind of instrument had been used. They felt that the true test of whether or not a person had achieved selected objectives on observational astronomy was to test that person over these objectives under the real sky. Therefore, they ran a limited study for which they developed an "Open-Sky" test with which to evaluate some objectives of planetarium education. The restricted study showed, in fact, that when evaluated with this instrument, the planetarium was superior to a classroom in which aids were limited to a chalkboard and celestial globe. They suggested further large-scale investigation into this problem.

The idea of using some sort of standardized open sky test to evaluate planetarium effectiveness is attractive. After all, instruction in astronomy is geared to objectives related to the understanding of the astronomical universe. Astronomy as a primarily observational science especially lends itself to an

understanding of many phenomena directly within the observational capabilities of the average citizen. Therefore, concepts that can be explored directly through observation should comprise a large portion of the pre-college curriculum and especially the astronomy portions of standardized science curricula. For evaluation of such instruction a well developed standardized open sky test would be invaluable. But there are some obvious problems with such an instrument. Use of an open sky test involves many inconveniences. Students must be brought together after school hours. Weather and sky conditions must be acceptable. Various celestial objects must be appropriately placed to aid in the administration of the instrument. For the most part such restrictions make open sky evaluation of astronomy curricula impractical.

On the other hand, the modern planetarium is used in an instructional and entertainment mode to simulate the visible sky from any place on earth at any time, past, present, or future, and completely independent of weather and sky conditions. In addition, students can visit the planetarium during school hours. If the planetarium were assumed to be an effective simulator of the open sky (a point not yet established by research), then to this investigator it seemed reasonable to suggest that perhaps a standardized planetarium test should have been developed as a reasonable and practical substitute for an open sky test. Thus, in this study the planetarium was used as such a substitute. The actual equivalence of the real sky and the planetarium is the so-called "transfer" problem and was not addressed in this study.

Several assumptions were made in the study:

1. The goal of all instruction in observational astronomy is increased familiarity with the real sky.

2. The planetarium is an effective simulator of the real sky.

3. The facilities used in developing this instrument were similar in the applicable aspects to typical facilities located at educational institutions. While the facilities at Fernbank Science Center are larger and more extensive (Zeiss Mark V, 70 foot dome, 500 seats) than planetariums typical of local schools, Fernbank is an integral part of the DeKalb (Georgia) County School System and is primarily concerned with the educational aspects of planetarium operation.

4. Each respondent's answers represented his own thinking and knowledge. Steps were taken to help insure this.

5. The jury technique is a valid method of determining the content validity of the instrument.

6. The interview technique is a valid method of determining the content validity of the instrument.

Development of the Instrument

Thirteen concepts on the annual motions of the sun, moon, and planets were developed by the investigator based upon the literature and experience. Since the concepts taught in various planetariums vary from place to place, there is no standard set of concepts established and accepted by the planetarium community from which to draw. Several criteria were required of each concept selected. First, each had to involve the annual motions of celestial bodies. This criteria was required simply to limit the scope of the content of the material to be covered. Thus, the PTAM in no way evaluates in general terms a representative sample of all planetarium instruction. The second criterion required each concept to describe a phenomenon that could be readily observed by students without auxiliary instruments.

Each concept had to be readily demonstrable in the planetarium and, finally, despite the fact that there is no generally accepted planetarium curriculum, it was required that the concepts generally be covered at the middle-elementary level. Concepts meeting these criteria were submitted to a jury of planetarium educators and professional astronomers for evaluation and validation.

Seventeen specific content objectives were written for the concepts. Each objective was written in behavioral terms and met another set of strict criteria. Each objective was required to measure that which the associated concept stated and be demonstrable in the planetarium. At least one test item was required to be devised for each objective and each had to state the behavior expected of students meeting the objective. The objectives were submitted to the same jury as validated the concepts, to evaluate the necessity and sufficiency of each objective. Necessity and sufficience were defined by positive answers to the following two questions:

1. Does each objective set forth a criterion <u>necessary</u> to demonstrate grasp of the associated concept?
2. Is the criterion set forth in each objective <u>sufficient</u> to establish grasp of the associated concept?

Finally two items were written to evaluate the attainment of each of the objectives. This resulted in a final instrument consisting of thirty-four items. Cirteria set for the items themselves included the necessity that they be demonstrable in the planetarium and involve a minimum of reading. Also they were required to be expressed in multiple-choice form. Each item had to be congruent with the task specified in the associated objective and sufficient to establish that that task had been accomplished.

These last two characteristics were validated for each item by the same jury described above.

The planetarium is unique as a teaching facility and, therefore, it would appear that the processes of science applied in the facility might be somewhat special when compared to the "normal" classroom. It also is conceivable that students operating in this facility might be called upon to use a particular set of intellectual skills in order to function effectively in the planetarium. Therefore, a process-skill matrix was developed to classify all objectives evaluated by the PTAM into those processes of science and intellectual skills that might be the most appropriate for sixth graders operating in a planetarium.

A second jury, this consisting of scientists and science educators, classified the objectives according to the processes of science used by sixth graders in meeting the criteria set forth in the objectives. A third jury, consisting of science educators, classified the objectives according to the intellectual skills elicited by each objective. These jury deliberations resulted in a three-dimensional matrix of processes, skills, and content, the first two dimensions of which are shown in Table 1. Subscales were then established for each process, skill, and content area using the matrix.

The revision of the PTAM into its final form was accomplished through a series of four pilot studies each using groups of sixth grade students from middle-class DeKalb County schools. While no items were discarded as a result of this process, some were revised primarily to increase acceptability of overly popular or unpopular distractors. The final pilot study using a school

Table 1
Final Process/Skill Classification of Objectives

	Intellectual Skill		
Process of Science	Recall	Discrimination	Concept Application
Using Space/Time Relationships	2, 7	12, 16	1, 8, 10, 11, 13, 14
Classifying	3, 6	15, 17	
Inferring	2, 4, 7	5, 12	1, 8, 9, 10, 11

Table 2
Criteria of Performance

Subscale	Number of Items	% Achievement Jolly Elementary	Criteria	
			No.	%
Using Space/Time Relationships	20	56.1	12	60.0
Classifying	8	58.8	5	62.5
Inferring	20	51.3	12	60.0
Recall	10	51.8	6	60.0
Discrimination	10	59.0	6	60.0
Concept Application	14	50.6	9	64.3
Each Content Objective	2		2	100.0

called Jolly Elementary was used to establish criteria for performance objectives on each of the subscales described above. These criteria are summarized in Table 2. The final version of the PTAM was then used in a field test involving 471 sixth grade DeKalb County students.

The Field Test

The chief problem of the field test was to determine whether sixth grade students would significantly improve their achievement in the areas of content, processes of science, and intellectual skills following 1) an hour-long single visit planetarium program or 2) an hour-long program in their school classroom over concepts of the annual motions of the sun, moon, and planets. These experimental treatments represent two of the formats of instruction in prevalent use at Fernbank Science Center. A second part of the problem for the field test was to determine whether there would be any significant differences in achievement in the subscale areas between the two groups.

The design of the field test was Campbell and Stanley's (1966) Experimental Design 6, the Posttest-Only Control Group Design. Since two treatments were used, there were two experimental and one control group. Table 3 illustrates the breakdown of the 471 students randomly assigned to the three groups:

Experimental Group Ep:	R	X	O	Planetarium instruction
Experimental Group Ec:	R	X'	O	Classroom instruction
Control Group C:	R		O	No instruction

The PTAM itself consists of thirty-four multiple-choice items demonstrable in the planetarium. Each student was provided an answer sheet, a pencil, and a masonite lapboard. Each question

Table 3
Subject Samples for the Field Test

School	Experimental Group		
	Ep	Ec	C
Heritage Elementary	21	21	22
Montclair Elementary	19	19	17
Rockbridge Elementary	36	42	40
Smoke Rise Elementary	40	43	44
Hambrick Elementary	36	35	36
TOTALS	152	160	159

was stated and choices pointed out in the planetarium sky. Most items inquire into the locations or motions of objects in the sky; a few require simple slides to be shown. Red cove lights were then brought up to enable the students to record their answers without loss of dark adaptation.

Data was obtained for each of the twenty-three subscales and the total scores. Table 4 is a summary of the subscore data. Criterion achievement scores were analyzed using an analysis of precentages method due to Garrett (1966) with significant (.05) between group differences being subjected to pair-wise comparisons using the Newman-Keuls Test. Mean achievement scores were analyzed using harmonic means analysis followed by pair-wise comparisons by the Newman-Keuls Test for significant (.05) differences. Table 5 summarizes significant differences found.

Several general trends or results were seen in this data. First, it appeared that both the planetarium and the school classroom treatments produced gains of a general overall nature more than specific mastery of selected concepts. There appeared to be no definite hierarchy of processes of science or intellectual skills under the circumstances in the study. Finally, the planetarium did fare better than the classroom on several content, process and skill scores.

Two final steps were taken to further establish the validity of the PTAM. The data from the field test was used in an item analysis and is included as Table 6. Also calculated were several indices of reliability appropriate to the design of the PTAM. Finally, an interview technique was used with a group of students

Table 4
Subscore Data

Score	% Meeting Criterion-Group Ep	Ec	C	Comb	Mean Raw Scores-Group Ep	Ec	C	Comb
Objective 1	25.7	23.1	15.1	21.2	1.02	.84	.71	.85
2	27.0	20.0	15.7	20.8	.98	.83	.73	.84
3	45.4	33.1	21.4	33.1	1.27	1.10	.85	1.07
4	14.5	7.5	8.2	10.0	.70	.59	.58	.62
5	23.7	25.0	27.7	25.5	.90	.99	.99	.96
6	32.9	41.9	28.3	34.4	1.15	1.27	1.03	1.16
7	63.2	33.1	18.2	37.8	1.55	1.06	.69	1.09
8	48.7	31.3	26.4	35.2	1.28	1.04	1.04	1.12
9	4.6	6.3	3.1	4.7	.45	.45	.50	.47
10	22.4	12.5	7.5	14.0	.91	.64	.63	.72
11	20.4	13.8	13.2	15.7	.91	.74	.71	.78
12	17.1	10.0	8.8	11.9	.87	.63	.57	.68
13	58.6	60.0	57.2	58.6	1.47	1.45	1.40	1.44
14	9.9	9.4	5.0	8.1	.44	.51	.38	.44
15	46.1	50.0	40.9	45.6	1.36	1.38	1.22	1.32
16	42.8	40.0	16.4	32.9	1.20	1.15	.77	1.04
17	50.7	50.6	21.4	40.8	1.38	1.41	.92	1.23
Using Space/Time Relationships	38.2	17.5	8.8	21.2	10.63	8.88	7.63	9.02
Classifying	67.4	67.5	40.9	58.6	5.17	5.16	4.02	4.78
Inferring	25.0	8.1	5.7	12.7	9.57	7.80	7.15	8.15
Recall	52.0	36.9	15.1	34.4	5.65	4.86	3.87	4.78
Discrimination	52.0	52.5	26.4	43.5	5.71	5.55	4.46	5.24
Concept Application	21.1	9.4	6.3	12.1	6.49	5.66	5.36	5.83
Total	32.9	21.9	7.5	20.6	17.85	16.07	13.72	15.85

Table 5
Summary of Significant Comparisons

Experimental Groups / Type of Score		Ep/C Mean	Ep/C Criterion	Ep/Ec Mean	Ep/Ec Criterion	Ec/C Mean	Ec/C Criterion
Hypothesis		1	2	3	4	5	6
Objective	1	Ep		Ep			
	2	Ep					
	3	Ep	Ep	Ep		Ec	
	4						
	5						
	6					Ec	
	7	Ep	Ep	Ep	Ep	Ec	Ec
	8	Ep	Ep	Ep	Ep		
	9						
	10	Ep	Ep	Ep			
	11	Ep		Ep			
	12	Ep		Ep			
	13						
	14						
	15						
	16	Ep	Ep			Ec	Ec
	17	Ep	Ep			Ec	Ec
Hypothesis		7	8	9	10	11	12
Using Space/Time Relationships		Ep	Ep	Ep	Ep	Ec	
Classifying		Ep	Ep			Ec	Ec
Inferring		Ep	Ep	Ep	Ep	Ec	
Hypothesis		13	14	15	16	17	18
Recall		Ep	Ep	Ep		Ec	Ec
Discrimination		Ep	Ep			Ec	Ec
Concept Application		Ep	Ep	Ep	Ep		
Hypothesis		19	20	21	22	23	24
Totals		Ep	Ep	Ep		Ec	Ec

Note 1: Groups favored are indicated

Table 6
Massed Item Analysis of the PTAM

Item No.	No. of Students A	B	C	D	Difficulty Index	Distractor Ratio	Point Biserial	Discrimination Index
1	76	186	136*	73	28.9	2.55	.076	.096
2	16	125	275*	52	58.4	7.81	.258	.242
3	110	51	247*	61	52.4	2.16	.438	.465
4	102*	159	66	143	21.7	2.41	.181	.096
5	61	111	234*	62	49.7	1.82	.252	.242
6	29	43	65	334*	70.9	2.24	.273	.287
7	56	162	107*	145	22.7	2.89	.282	.197
8	77	85	259*	49	55.0	1.73	.271	.274
9	57	68	62	283*	60.1	1.19	.293	.331
10	191*	203	56	21	40.6	9.67	.240	.223
11	76	221*	120	53	46.9	2.26	.295	.338
12	59	253*	109	49	53.7	2.22	.217	.242
13	38	342*	49	41	72.6	1.29	.400	.420
14	250*	83	29	109	53.1	3.76	.426	.478
15	96	153	118*	104	25.1	1.59	.016	.000
16	80	65	267*	58	56.7	1.38	.410	.452
17	204	104	70	92*	19.5	2.91	.175	.146
18	58	77	114	220*	46.7	1.97	.379	.471
19	333*	41	36	59	70.7	1.64	.355	.357
20	90	82	117	182*	38.6	1.43	.400	.465
21	79	207*	137	48	43.9	2.85	.325	.385
22	91	210*	75	93	44.6	1.24	.264	.261
23	249*	72	39	109	52.9	2.79	.399	.459
24	77	77	87	230*	48.8	1.13	.419	.433
25	58	93	121	195*	41.4	2.09	.322	.344
26	164	105	82	118*	25.1	2.00	.208	.153
27	84	123	191*	73	40.6	1.68	.280	.293
28	39	138	200*	94	42.5	3.54	.141	.127
29	116	158*	107	86	33.5	1.35	.165	.223
30	36	50	335*	50	71.1	1.39	.284	.255
31	74	135	175*	87	37.2	3.97	.347	.369
32	143	132*	127	69	28.0	2.07	.267	.236
33	16	60	32	362*	76.9	3.75	.280	.280
34	259*	55	64	92	55.0	1.67	.353	.401

*Indicates the correct choice

Kuder-Richardson Formula 20 Reliability = .66
Split Half Reliability = .71
Livingston Criterion-Referenced
Reliability for Criterion Score
of 20 = .85

in order to validate the PTAM against the students' grasps of the concepts involved.

Summary

A reliable testing instrument, the Planetarium Test on Annual Motions (PTAM), that uses the planetarium in an observational mode as a testing facility was developed. This instrument can be used to evaluate any instruction whose objectives are the same as those evaluated by the instrument. The PTAM was used to evaluate and compare the results of two forms of instructional programs at Fernbank Science Center. In general, the planetarium was found to fare better than the classroom in producing gains in achievement in several content, processes of science, and intellectual skills areas.

SELECTED BIBLIOGRAPHY

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago, IL: Rand McNally & Company, 1966.

Dean, N., & Lauck, G. M. Planetarium instruction - using an open-sky test. Science Teacher, 1972, 39, 54-55.

Garrett, H. E. Statistics in psychology and education (6th ed.). New York, NY: David McKay Company, Inc., 1966.

Livingston, S. A. Criterion-referenced applications of classical test theory. Journal of Educational Measurement, 1972, 9, 139-140.

Warneking, G. E. Planetarium education in the 1970's - a time for assessment. Science Teacher, 1970, 37, 14-15.